MCI

MCI Telecommunications Corporation

1801 Pennsylvania Avenue, NW
Washington, DC 20006
202 887 2380
FAX 202 887 3175
VNET 220 2380
2181493@MCIMAIL.COM
MCI Mail ID 218-1493

Karen T. Reidy
Attorney
Federal Law and Public Policy

ORIGINAL

EX PARTE OR LATE FILED

ORIGINAL

SEP 1999

September 7, 1999

VIA HAND DELIVERY

Ms. Magalie Roman Salas, Secretary
Federal Communications Commission
Office of the Secretary - Room TWB-204
445 Twelfth Street, SW
Washington, DC 20554

Re: Ex Parte: CC Docket Nos. 98-121 and 98-56

Dear Ms. Salas:

On September 3, 1999, I sent the attached document to Daniel Shiman of the Common Carrier Bureau's Policy and Program Planning Division. Please include this filing in the record of the above-referenced proceedings.

Two copies of this Notice are being submitted in accordance with Section 1.1206 of the Commission's rules.

Sincerely,

Karen Reidy

Karen T. Reidy

Attachment

cc: Daniel Shiman (w/attachment)

No. of Copies rec'd 0+3
List ABCDE

Applying Common Means Tests to Determine Parity Performance in the Provision of Benchmark Service Quality Measures

JOHN D. JACKSON, *Auburn University, 203 Lowder Business Building, Auburn, AL 36849.*

The Telecommunications Act of 1996 has spawned an entire body of research relating to service quality parity. In the main, this research addresses the problem of determining whether the ILEC provides the same quality of service to a given CLEC that it provides to itself. There are two types of service quality measures that must be considered: (1) *Analogue measures* apply to services that the ILEC provides to its customers as well as to the CLECs (i.e., the ILEC's service is strictly analogous to that provided to the CLEC). (2) *Benchmark measures* apply to services produced by the ILEC for the CLECs but not for its own customers. [1]

For Analogue measures, parity requires not only that the CLEC and ILEC means be equal but also that their variances be equal. Many analysts miss this second requirement. Intuitively, the CLEC would be put at a competitive disadvantage even with the same average level of service provision as the ILEC, if in addition, the CLEC's service was more dispersed (more uneven) than the ILEC's. This leads to a set of means difference tests, one of which has more power than the others in detecting departures from the equal variance requirement.

Benchmark measures, on the other hand have no sample of ILEC observations with which they can be compared. Rather the CLEC data must be compared to a set standard or benchmark determined by a panel of experts from the CLECs and the ILEC. As such they differ fundamentally from analogue measures, and the relevant statistical tests differ as well. Here parity requires that

---

[1] CLECs have proposed that a benchmark standard should also be established for services with retail analogues, as a minimum standard, in addition to a rolling parity standard. Fixed standards (rather than rolling parity standards) are needed to allow CLECs to plan internal processes and operations and to allow CLECs to provide dependable dates and time periods to their customers. In such cases, the observations about benchmark measurements and statistical testing noted below still apply as the benchmark generally is set after taking into account random variations in analogue processes.

the CLEC mean not differ from the value prescribed in the benchmark. Since the prescribed benchmark has no variance, and since only CLEC data enter the calculation of the appropriate statistic, there is no corresponding variance equality requirement.

Testing Analogue Measures

A consensus seems to have arisen among the various states as to the appropriateness of the LCUG Z for parity tests involving these measures, at least for large samples. The rationale is presented below, as briefly as completeness will allow: The Central Limit Theorem and a theorem from statistical distribution theory allow us to assert that ($\overline{X}_{ILEC}$ - $\overline{X}_{CLEC}$) follows a normal distribution with mean ($\mu_{ILEC} - \mu_{CLEC}$) and variance $[(\sigma^2_{ILEC}/n_{ILEC}) + (\sigma^2_{CLEC}/n_{CLEC})]$, so that the random variable

$$Z = \frac{(\overline{X}_{CLEC} - \overline{X}_{ILEC}) - (\mu_{CLEC} - \mu_{ILEC})}{\sqrt{\left(\frac{\sigma^2_{CLEC}}{n_{CLEC}} + \frac{\sigma^2_{ILEC}}{n_{ILEC}}\right)}}. \tag{1}$$

follows a standard normal distribution for large samples. Z forms the basis for analogue parity testing. Recall that for analogue measures, the parity question requires the test of: H_0: $\mu_{ILEC} = \mu_{CLEC}$ *and* $\sigma^2_{ILEC} = \sigma^2_{CLEC}$.

Substituting these constraints into equation (1), we have

$$Z = \frac{(\overline{X}_{CLEC} - \overline{X}_{ILEC})}{\sigma\sqrt{\left(\frac{1}{n_{CLEC}} + \frac{1}{n_{ILEC}}\right)}} \tag{2}$$

where σ is the square root of the common variance σ^2 (= $\sigma^2_{ILEC} = \sigma^2_{CLEC}$). Next we note that the random variable ($\phi S^2/\sigma^2$) follows a χ^2 distribution with ϕ degrees of freedom. Taking the square root of this expression divided by its degrees of freedom and dividing the result into (2), yields

$$Z = \frac{(\overline{X}_{CLEC} - \overline{X}_{ILEC})}{S\sqrt{\left(\frac{1}{n_{CLEC}} + \frac{1}{n_{ILEC}}\right)}} \tag{3}$$

Technically, this statistic follows a Student's t distribution whose degrees of freedom (ϕ) are inherited from S. However, since the statistic is valid only for large samples and since the distinction between the t distribution and the standard normal (Z) vanishes with increased sample size, equation (3) is sometimes termed a Z statistic. We will follow this convention.

We refer to equation (3) as a "parent statistic" because it can give rise to many forms, depending on how one chooses to estimate S. Two of these forms, one using a pooled variance estimator (Z_{pooled}) and one using the ILEC variance estimator (Z_{LCUG}), are particularly relevant to the problem of testing for parity. Before defining and comparing these two statistics, however, it is important to note that one commonly used form of (3) is not suitable for parity testing.

Often, one sees a version of (3), which amounts to substituting the sample variances, S^2_{ILEC} and S^2_{CLEC}, for their corresponding parameters in (1). This statistic is used to test means differences when the population variances are unequal. It follows a t distribution and requires a "degrees of freedom adjustment" to be accurate. Since parity requires the variances to be equal and since this statistic would be appropriate only when $\sigma^2_{ILEC} \neq \sigma^2_{CLEC}$, it is particularly ill suited to testing for parity. Alternatively, the two variants of (3) discussed below are both well suited to parity testing, however, one provides a more powerful test of the parity hypothesis than the other.

Since our objective is to find an appropriate estimate S^2 of the common variance σ^2 to be used in (3), an obvious procedure is to simply take a weighted average of the CLEC and ILEC variance estimates. If we take the weights to be the percent of the total degrees of freedom attributable to each carrier, we obtain the traditional pooled variance estimator

$$S^2_{pooled} = \frac{(n_{CLEC} - 1)S^2_{CLEC} + (n_{ILEC} - 1)S^2_{ILEC}}{n_{CLEC} + n_{ILEC} - 2}$$

where the CLEC weight is $[(n_{CLEC} - 1)/(n_{CLEC} + n_{ILEC} - 2)]$, the ILEC weight is $[(n_{ILEC} - 1)/(n_{CLEC} + n_{ILEC} - 2)]$, and the total degrees of freedom is $(n_{CLEC} + n_{ILEC} - 2)$. It is worth noting that since the weights are normalized, i.e., they sum to one, the value of S^2_{pooled} will always lie between the values of S^2_{CLEC} and S^2_{ILEC}. Substituting S_p ($= \sqrt{S^2_p}$) into (3) yields

$$Z_{pooled} = \frac{(\overline{X}_{CLEC} - \overline{X}_{ILEC})}{S_p\sqrt{\left(\frac{1}{n_{CLEC}} + \frac{1}{n_{ILEC}}\right)}} \tag{4}$$

which we shall refer to as the pooled Z. It follows a standard normal distribution for large samples.

An alternative to the traditional approach of using Z_{pooled} was proposed by LCUG in February 1998; the LCUG document describing the test in detail is attached. This approach amounts to substituting the estimated ILEC standard deviation S_{ILEC} ($=\sqrt{S^2_{ILEC}}$) for S in (3). At first glance this approach might seem overly simplistic, but it turns out to have substantial intuitive appeal, and it produces a statistically more powerful test of the parity hypothesis than the traditional approach -- as will be demonstrated below. The test statistic, which we shall term the LCUG Z (or Z_{LCUG}), can be seen to be

$$Z_{LCUG} = \frac{(\overline{X}_{CLEC} - \overline{X}_{ILEC})}{S_{ILEC}\sqrt{\left(\frac{1}{n_{CLEC}} + \frac{1}{n_{ILEC}}\right)}} \tag{5}$$

While both Z_{pooled} and Z_{LCUG} have power to detect violations in the form of means differences (since both statistics have the same numerator), the LCUG Z is a more powerful test of parity since it also incorporates an indirect test of equality of varianсs.

To see this, first note that the traditional method requires two tests to establish parity or lack thereof. A test of H_0: $\mu_{CLEC} = \mu_{ILEC}$ using Z_{pooled} must be coupled with a test of H_0: $\sigma_{CLEC} = \sigma_{ILEC}$. This second test typically employs an F statistic computed as the ratio of the CLEC and ILEC variance estimates. Parity requires that neither null be rejected. It is important to note that the necessity of using two tests to investigate parity reduces the power of each test. Thus we would prefer a single test that can detect violations of parity due to both differences in means and differences in variances. The LCUG Z provides such a test statistic.

From the standpoint of comparing variances, the fact that the ILEC is required to provide the CLEC with *at least the same* service level means that the ILEC variance is the relevant standard of comparison. If the CLEC variance exceeds the ILEC variance, then parity in service cannot be accepted. Moreover, since the ILEC samples are typically quite large (many times, in the hundreds of thousands), they may be expected to provide very accurate estimates of the variances in the relevant ILEC performance measures.

For a given means difference, a more powerful test of parity would be more likely to reject the null if $\sigma^2_{CLEC} > \sigma^2_{ILEC}$ and less likely to reject if $\sigma^2_{CLEC} < \sigma^2_{ILEC}$. If $\sigma^2_{CLEC} > \sigma^2_{ILEC}$ we would expect $S^2_{CLEC} > S^2_{ILEC}$ so that $S^2_p > S^2_{ILEC}$ and hence $|Z_{LCUG}| > |Z_{pooled}|$. Thus when parity is not present, Z_{LCUG} would be more likely to reject the null than Z_{pooled}. On the other hand, if $\sigma^2_{CLEC} < \sigma^2_{ILEC}$ we would expect $S^2_{CLEC} < S^2_{ILEC}$ so that $S^2_p < S^2_{ILEC}$ and hence $|Z_{LCUG}| < |Z_{pooled}|$. Thus when parity is present, Z_{LCUG} would be less likely to reject the null than Z_{pooled}. Of course, when equality holds the two approaches produce identical results. It should now be clear why the LCUG Z provides a more powerful test of the parity hypothesis. *It is this result – that a violation of parity occurs when $\sigma^2_{CLEC} > \sigma^2_{ILEC}$ and that the LCUG Z will reject more often than other tests when $\sigma^2_{CLEC} > \sigma^2_{ILEC}$ – that gives the LCUG Z its superiority over its competing tests.*

Testing Bootstrap Measures

Statistical tests should ***not*** be applied to benchmark measures to evaluate parity in service. In order to understand this, we must first look at the ILEC argument for statistical testing, which can be summarized as follows: Parity is empirically investigated in the provision of benchmark services using a sample of CLEC data. Because of sampling variability, process variability, or random variation, it is possible for the sample mean to be less than the benchmark value even though the true population mean/process mean is equal to or exceeds the benchmark value. Thus we must test to see if the sample mean is far enough below the benchmark value so that the difference is not attributable to chance; i.e., we must test whether the population/process mean is statistically significantly less than the benchmark. This requires a test of the null hypothesis H_0: $\mu_{CLEC} = \mu_0$, where μ_0 is the prescribed benchmark value. The statistical underpinnings of the appropriate test are given below.

Analyzing only a sample of n_{CLEC} observations, the Central Limit theorem tells us that for large n_{CLEC} , $\overline{X}_{CLEC}$ follows a normal distribution with mean μ_{CLEC} and variance σ^2_{CLEC}/n_{CLEC} , so that the random variable

$$Z = \frac{\overline{X}_{CLEC} - \mu_0}{\frac{\sigma_{CLEC}}{\sqrt{n_{CLEC}}}} \qquad (6)$$

follows a standard normal distribution. Unfortunately, we cannot operationalize this result because we do not know σ^2_{CLEC}. However, we know, also based on the Central Limit Theorem, that the random variable

$$\chi^2 = \frac{(n_{CLEC} - 1)S^2_{CLEC}}{\sigma^2_{CLEC}} \tag{7}$$

follows a χ^2 distribution with (n_{CLEC} - 1) degrees of freedom. Divide equation (7) by its degrees of freedom, take the square root of the result and divide it into (6) to obtain

$$Z_B = \frac{\overline{X}_{CLEC} - \mu_0}{\frac{S_{CLEC}}{\sqrt{n_{CLEC}}}}. \tag{8}$$

The random variable Z_B follows a t distribution with (n_{CLEC} - 1) degrees of freedom. Since it is valid only for large samples, it is often said to simply follow a Z (standard normal) distribution. Equation (8) is operational under H_0, that is, it can be computed based on sample information assuming that the null is true. Therefore, this statistic can legitimately be used to test hypotheses of the form H_0: $\mu_{CLEC} = \mu_0$.

A number of alternatives to equation (8) have been proposed. Among the more interesting ones were (a) replacing S^2_{CLEC} with S^2_{ILEC} (presumably based on some appeal to the LCUG Z) and (b) replacing S^2_{CLEC} by 1 (which was proposed by the Texas PUC staff). Regarding (a), if S^2_{ILEC} does in fact exist, we are using the wrong test: the measure is appropriately an analogue measure and the LCUG Z is the appropriate test[2]. Disregarding the *ad hocery* of (b) momentarily, it should be clear that substituting any value for S^2_{ILEC} destroys the statistical underpinnings of the test. That is, the resulting statistic does not follow a standard normal distribution, which the test requires. In summary, of all of the (large sample) tests on the table for testing H_0: $\mu_{CLEC} = \mu_0$, equation (8) is the only one that is statistically appropriate.

The problem with basing a parity test on equation (8) is not statistical, it is economic. The rationale is simple and inescapable. A given level of significance

---

[2] This criticism, of course, is moot if the benchmark measure also has a retail analogue (see footnote 1).

(α) establishes the critical value of the test statistic: $\alpha = .15$ indicates a critical value of $Z = 1.04$, $\alpha = .10$ implies a critical $Z = 1.28$, and α=0.05 gives a critical $Z = 1.645$ for one tailed tests. Values of Z_B in excess of these critical values indicate rejection of parity. Thus, whatever the prescribed critical value of Z, the smaller the value of Z_B, the less likely we are to reject the null hypothesis of parity (i.e., H_0: $\mu_{CLEC} = \mu_0$), *ceteris paribus*. It remains to simply note that the larger S_{CLEC}, the smaller Z_B. This means that those who advocate the LCUG Z to determine parity of analogue measures, because it incents the ILEC to provide the CLECs less variable service, could not logically advocate this test for benchmark measures since it incents the ILEC to do exactly the opposite, i.e., provide more variable service.

Presumably the ILEC wishes to avoid being penalized for providing poor service. As far as benchmarks are concerned, one way to accomplish this end is to decrease the relevant value of Z_B, and one way to do that is to ***increase*** the value of S_{CLEC}. *Thus the ILEC has an incentive to provide inferior service to reduce its penalty*[3]. This is a perverse outcome and presumably not one that the Telecommunications Act of 1996 would have anticipated or desired; it ***must*** be avoided at all costs. In summary, *it follows that statistical testing of Benchmark Measures is not appropriate, since the only legitimate statistical test provides the perverse incentive to discriminate against the CLEC through increased variability of service in order to pass the test.*

How, then, can we account for the various sources of variation inherent in the CLEC sample? We contend that they have already been taken into account in the establishment of the benchmark values. These values do not demand perfection; the experts who negotiated them set them low enough to allow for process variability. If the ILECs now think these values too high, they should seek changes in the tariffs, interconnection agreements and regulatory orders that established them--often long before the idea of a statistical test of parity was proposed. In any case, the answer to the sampling variability problem for benchmark measures is not statistical testing. ILECs merely have seized this as an opportunity to gain some additional leeway in discriminating against their competitors without suffering financial consequences.

[3] See, Nicholas Economides, "The Incentive for Non-Price Discrimination by an Input Monopolist," *International Journal of Industrial Organization*, 16:271-84. (May 1998).